Childhood characteristics and participation in Scottish Mental Survey 1947 6-day sample follow-ups

Childhood Characteristics and Participation in Scottish Mental Survey 1947 6-Day Sample Follow-Ups: Implications for Participation in Aging Studies

Wendy Johnson[1], Caroline E. Brett, Catherine Calvin[1] and Ian J. Deary[1]

[1]MRC Centre for Cognitive Ageing and Cognitive Epidemiology, Department of Psychology, University of Edinburgh, UK

Correspondence to: Wendy Johnson, Centre for Cognitive Ageing and Cognitive Epidemiology and Department of Psychology, 7 George Square, University of Edinburgh, Edinburgh, UK EH8 9JZ, phone +44(0)131 651 1304, fax +44(0)131 650 3461, email wendy.johnson@ed.ac.uk

Keywords: Cognitive ability, personality, study participation, childhood, young adulthood, old age, aging studies

Abstract

Given the ‘graying’ of especially the populations of most western nations, studies of factors contributing to well-being in later life are important and common. It is important to their accuracy that they be based on samples representative of the populations in the relevant age groups. There is general awareness that several characteristics such as sex, socioeconomic status, cognitive ability and personality are associated with study participation, but many researchers assume that this reflects life circumstances at time of recruitment rather than inherent individual characteristics that shape those circumstances throughout people’s lives. The Scottish Mental Survey 1947 6-Day Sample Follow-Up Study offered an unusual opportunity to test this assumption, as follow-up study participation data were available both in young adulthood and at age 77. Participation at age 77 was dramatically restricted relative to that in young adulthood. Cognitive abilities and a composite of conscientiousness-related variables independent of cognitive ability assessed in childhood predicted participation at young ages, but much more strongly at older ages. Evidence was available that these results were not specific to the recruiting and assessment methods used in this study. This suggests that participation in studies of aging is a function not just of contemporaneous circumstances but also of early-life cognitive and personality characteristics that have shaped those circumstances. .

The populations of most nations are ‘graying’ rapidly, especially those of western nations. That is, the proportion of the world’s population that can be considered in old age is growing rapidly (United Nations, 2009). Because declines in physical health, well-being, and cognitive function in old age are common, there is great interest in identifying factors contributing to healthy aging, and studies in this area are frequent and ongoing. It is important to their accuracy that such studies be based on samples that are representative of the populations in the relevant age groups (Menard, 2002; Taris, 2000). There is general awareness that several characteristics such as sex, socioeconomic status, cognitive ability and personality are associated with study participation (Menard, 2002; Taris, 2000), but many researchers assume that this reflects life circumstances at time of recruitment rather than inherent individual characteristics that shape those circumstances throughout people’s lives (Nishiwaki, Clark, Morton, & Leon, 2005). Many also assume that similar characteristics are involved in participation across age groups, and affect it to similar degrees.

The question of sample selectivity, however, has attracted increasing attention (e.g., Hunt & Madhyastha, 2008; Murray, Johnson, McGue, & Iacono, 2014; Rabbitt, Lunn, & Wong, 2008), especially in epidemiological studies. These studies tend to focus on participation rates, in hopes that, if recruitment has been broad and participation rate high, the sample will be population-representative. Population-representativeness can be partially verified, when relevant population registry data are available, and/or some random group of non-participants can be persuaded to complete an abbreviated version of the assessment (Stang, 2003). These solutions are not ideal, however, as population registries rarely compile information on the kinds of individual psychological characteristics that cause people to balk at participating, and those who are willing to complete an abbreviated assessment likely differ from those who are not. Another approach is to make statistical adjustments to study estimates using sensitivity analyses (e.g., de Luna & Lundin, 2014), simulations of potential bias in participation (e.g., Roth et al., 2014), and/or techniques such as propensity score matching (e.g., Boutwell, Beaver, & Barnes, 2012). All of these techniques require quantitative assumptions, however, and it is often difficult to evaluate the extent to which these assumptions may hold.

The potential for sample selection to create bias is especially acute in longitudinal studies because whatever characteristics affect participation rates at time of recruitment tend also to affect willingness to complete follow-up assessments (Menard, 2002; Stang, 2003; Taris, 2000). This

causes samples to become increasingly more selective over time. Often this means increasing selection on the characteristics that drove initial participation, but it can also mean introduction of new selection characteristics when there is sufficient time between assessments that systematic changes in participants' life circumstances take place. For example, at age 52, when most people are still actively employed and may be at the peaks of their careers, with offspring still at home or being supported at university, the factors that drive willingness to participate in a research study may be related to competing professional and personal time demands that do not apply 15 years later, when many would have retired and most offspring likely achieved independence, but health limitations may have begun to have relevance. This can be especially problematic when samples span large age ranges at each assessment (Sliwinski, Hoffman, & Hofer, 2010).

Tendency for people with higher cognitive abilities to be more likely to participate in research studies is one of the sample selection processes that has been studied most extensively. This makes the studies in this area a good sample of the state of the art. Most studies that have addressed whether study samples were selected for higher cognitive ability have found evidence that they were, and that the selection became more extreme with each assessment wave in longitudinal studies (e.g., Beaver, 2013; Cooney, Schaie, & Willis, 1998; Dykiert, Gale, & Deary, 2009; Nishiwaki, Clark, Morton, & Leon, 2005). Consistency of results has not been uniform, however (e.g., Kerr, Lambert, & Bem, 1996; Lynam, Moffitt, & Stouthhammer-Loeber, 1993), and degrees of selection tend to vary with participant age, in general becoming more extreme with increasing age. There are at least two reasons for this (Johnson, McGue, & Deary, 2013). First, people with greater cognitive ability tend to be more interested in and more likely to appreciate the importance of scientific research, and, second, people with greater cognitive ability tend to be more likely to survive in good health from one age to another, increasingly so with greater age (Calvin, et al., 2011). Samples that can be used to evaluate the extent to which this phenomenon reflects stable cognitive abilities across the lifespan and compare participation rates in different phases of adulthood, however, are rare, as are samples that can be used to examine other psychological characteristics for analogous properties. Moreover, few studies can be considered truly population-representative even when initial recruitment takes place in childhood, so bases of comparison can generally only be relative. The purpose of this study was to explore participation selection effects of cognitive ability, personality, and other personal

characteristics in an unusually population-representative sample initially surveyed in childhood, followed through young adulthood from ages 14 to 27, and then re-recruited at age 77.

METHOD

Participants – The 6-Day Sample and Follow-Up Studies

In 1932 and 1947, the Scottish Council for Research in Education (SCRE) conducted Mental Surveys of almost all children born in 1921 and 1936, respectively, who were attending schools in Scotland (Scottish Council for Research in Education, 1933; 1949). The purpose of the first survey in 1932 (n=87-498) was to observe the population distribution of cognitive ability; the second in 1947 (n=70,805) was conducted to examine to what degree the population distribution might have changed. Extensive efforts were made to test all schoolchildren in Scotland born in the targeted years, even when they were in remedial or special education programs or suffered other disabilities. This makes these two samples among the most completely population-representative ever. A representative sample of participants in this second survey, born on the first three days of each month of 1936 (thus effectively randomly selected with respect to any variable related to cognitive ability), and their families and teachers completed a more extensive Sociological Survey at age 14 in 1950. Participants born on the first day of any even-numbered month received an additional cognitive assessment and were re-assessed annually on a number of factors from ages 15-27 (1951-1963). These 1208 (618 female) participants were called the 6-Day Sample (MacPherson, 1958; Maxwell, 1969). The additional assessments were administered by teachers, educational psychologists, and survey administrators representing SCRE, who visited and surveyed participants' homes and interviewed their parents and sample members themselves, and by head teachers while participants were at school. The assessments included further psychological measures of intelligence and personality, details of socioeconomic circumstances, school attendance, and, after participants left secondary school details of further schooling, employment, marriage and family.

In 2012, the original 6-Day sample participants were traced through United Kingdom and Scottish population records, recording deaths and their causes and locating as many of those surviving as possible (Brett & Deary, in press). In late 2012 and 2013, located participants residing in Scotland, England, or Wales received a postal recruitment invitation explaining a follow-up study to explore health status, demographic circumstances, psychological characteristics, well-being and attitudes toward life at age 77, and their associations with the data collected in the earlier

assessments. The invitation included a self-administered assessment package. Participants were requested to return a one-page form indicating willingness (or not) to participate, and to return the assessment package by mail when it was completed.

Of the original 1208 participants, 417 were deceased (164 females), 68 could not be located, and 89 had emigrated from the United Kingdom. The remaining 635 (including 1 earlier emigrant; 370 females) were invited to participate; 1 had emigrated, 2 were deceased, and 20 were deemed not capable by English/Welsh law since they had been located. No replies were received from 205 despite follow-up mailing, 138 refused participation, and 205 indicated willingness to participate, either by completing the one-page form or telephoning the study office. The primary reason for refusal was lack of interest. Completed assessments were received from 171 (90 females), for a participation rate of 27% of those invited. This participation rate may sound low, but the efforts involved in locating or accounting for all the original participants of the unusually population-representative 6-Day Sample were extensive, including matching to National Health Service medical records.

Measures

Participations in the 13 follow-up young-adult assessments to the original 6-Day Sample Study (*n* = 1208) following the Scottish Mental Survey of 1947 (SMS1947) from 1951 through 1963, and the age-77 Follow-Up Study (*n* = 171) was were the primary outcomes we analysed. We studied the associations between participation patterns in these subsequent follow-up waves and relevant variables from the original 6-Day Sample childhood assessment, including the following:

Moray House Test #12 (MHT). Most (1112) of the 6-Day Sample participants completed the MHT on June 4, 1947 as part of the SMS1947, when they were age 11. The MHT (Scottish Council for Research in Education, 1933) is a valid, group-administered test of cognitive ability, requiring 45 minutes to administer. It consists of 71 items. Verbal reasoning items predominate, but there are, some numerical and other types of items. It generates a maximum score of 76. Of the 75,211 born in 1936, 70,805 received the test; those who did not complete it were not in attendance at school on the day of administration. SMS1947 was thus effectively an assessment of cognitive ability in the whole relevant population.

Terman-Merrill IQ Test – 1937 Revision (TMIQ). Participants in the 6-Day Sample also completed the individually-administered L form of the TMIQ (Scottish Council for Research in

Education, 1949). One of the purposes of administering this test was to corroborate validity of the MHT; the correlation between the two tests' scores was .80.

Father's Social Class. Father's highest attained occupation was recorded during the sociological assessment. This was coded by the age-77 Follow-Up Study research team according to the five-class system of the1951 Classification of Occupations used in the 1951 Population Census (Knight, 1967). In this system, Class I is professional, II intermediate, III both manual and non-manual skilled, IV semi-skilled, and V unskilled. This is not the classification system that was used by SCRE in its official publications, but it was more commonly used in official publications of the period, and in research relating to education and social mobility (e.g., Johnson, Brett, & Deary, 2010; Paterson, Pattie, & Deary, 2011).

Height was measured in inches at age 11. We considered height because it is commonly considered to reflect social class and/or general constitutional robustness, and samples are often healthier than their underlying populations.

Personality. Participants' head teachers assessed six areas of their personalities as part of the First School Schedule in 1950. They rated self-confidence, perseverance, mood stability, conscientiousness, originality, and desire to excel on a 5-point scale, with 1 referring to 'marked lack', 2 'less than average', 3 'average', 4 'more than average', and 5 'very'. All these personality characteristics tend to be related conceptually and empirically to school achievement, as are cognitive ability scores, and teachers tend to observe pupils primarily within the achievement-related school setting. Thus it is not surprising that the personality ratings correlated substantially with the cognitive ability test scores (correlations ranged from .19 to .43). For purposes of this study, we thus regressed TMIQ scores from the personality ratings, making use of the intelligence-adjusted residuals as our measures of personality.

Tabulation of Data from 1951-1963 Follow-Up Assessments. During the course of the follow-up assessments, participants reported whether they had changed residence; changes in job; health problems encountered; and incidence of personal problems such as marital distress or separation, unemployment, trouble with the law, miscarriage, and illness or death of a child. We summed these reports over time, divided by number of completed assessments to adjust for tendency to obtain information at all, and standardized the results to produce measures of these incidences that could be considered comparable among participants who participated to different degrees in the follow-up

assessments. Participants also reported whether they were in tertiary education programs at each assessment. We summed these and divided by the number of completed assessments to obtain an analogous relative measure of years in school.

ANALYSIS AND RESULTS

Population-Representativeness of the 6-Day Sample

The process used to recruit the 6-Day Sample was arbitrary, based on birth date on the first of an even-numbered month in 1936. It was intended to generate a random subsample of the SMS1947 cohort. We examined the extent to which this appeared to have been the case, using the limited data available on the SMS1947. Results are shown in Table 1. Of course there were mean differences and all were significant, given the large size of the SMS1947 cohort. But most of the effect sizes of the differences were trivial, and even standard deviations and skews were highly similar. The 6-Day sample was slightly older than the full cohort, but the difference corresponded to 18 days. This is about the average difference between the first and last day of any month, and thus almost exactly what would be expected based on the method of recruitment. The 6-Day Sample did have a slightly higher proportion of females than did SMS1947 (less than 2% difference).

Participation Patterns in the 1951-1963 Follow-Up Assessments

Participation rates in the 1951-1963 follow-up assessments were high. This was due at least in part to the extensive retention efforts made by the survey administrators representing SCRE. It may also have been due in part to the post-World War II need to rebuild much of the British economy and infrastructure, which may have kept sample members relatively local and thus easier to trace from year to year. Fully 53% of the sample completed all 13 assessments, 16% completed all but 1, and another 13% completed all but 2, for a total of 72% missing no more than 2. The 1952 and 1953 assessments were the ones most commonly missed. These were conducted at ages 16 and 17, a time when many participants were entering the job market, military service, or university and likely leaving the homes in which they had grown up, perhaps making them harder to locate. Another possibility is that missing these assessments had nothing to do with participants, as it appears that the home visitors in some locations may not have administered assessments in one or both of these years at all. The average participation rate was 87%. Full participation data are summarized in Table 2.

We used latent class analysis to examine participation patterns in the 1951-1963 follow-up assessments. This technique relies on a measurement model similar to that of factor analysis, but produces groups of individuals analogous to those from cluster analysis, rather than the groups of variables produced by factor analysis. The method relies on the presence of 'breakpoints' or gaps in data distributions that presumably identify qualitatively and categorically distinct subgroups within a population. We did not, however, actually presume that such qualitatively and categorically distinct groupings of assessment participation exist, in general among young adults, or in the specific case of the 6-Day Sample. Rather, we suspected that individual differences in propensity to participate in research studies, and the resulting patterns, are continuous in nature, and that any breakpoints that could be established to group those patterns would be somewhat arbitrary. As Nagin and Tremblay (2005) discussed with respect to groups of developmental trajectories, we applied the technique because of its power to provide descriptive information that could be linked to other participant characteristics that might illuminate factors involved in study participation. This theoretical orientation offered two important technical advantages.

First, as in factor analysis or cluster analysis, the number of latent classes extracted strongly influences the nature of results. Because we explicitly rejected the presumption that latent participation taxons exist, we also rejected the idea that a primary goal was to identify the specific, immutable, theoretically meaningful number of such participation groups in the population at large, or even in our sample. Any solution for which the indicated participation groups offered interpretive descriptive power would be valuable. Researchers hoping to determine one specific optimal or true number of taxons generally rely on several of the many available fit statistics, but, as often noted, (e.g., Bauer & Curran, 2003; Bentler, 1990; Markon & Krueger, 2004), these indices do not always clearly indicate a single best-fitting model, and there is a tendency for fit indices to improve continuously with greater numbers of classes. The optimal balance in the trade-off between optimal fit and parsimony is often more a matter of judgment than objective observation. We were seeking a practical descriptive tool rather than naturally occurring breakpoints between pre-existing taxons. Thus, inability to identify a clearly optimal number of groups objectively did not present any conceptual difficulty as long as the number of groups selected offered some descriptive advantage. We evaluated model fit using the Akaike Information Criterion (AIC; Akaike, 1983), the Bayesian Information Criterion (BIC; Raftery, 1995) and the sample-size-adjusted BIC (adjBIC; Sclove, 1987).

Smaller values indicate better model fit and favor more parsimonious models. We supplemented these with two tests that indicate whether a model with one less group would fit better, the Parametric Bootstrapped (McLachlan & Peel, 2000) and Lo-Mendell-Rubin Adjusted (Lo, Mendell, & Rubin, 2001) Likelihood Ratio Tests, as well as other checks that the selected solution offered the (admittedly subjectively judged) optimal interpretation of the data.

Second, as Bauer and Curran (2003) have pointed out, group modeling using maximum likelihood estimation relies on the assumption that data are normally distributed. For the categorical participation yes/no data analysed here, this means assumption of a normally distributed continuous potential for participation, coupled with a threshold for actual participation in the 6-Day assessments. Variations in apparent threshold across assessments or deviations from normality in any such actual potential could give the appearance of taxons in mixture models of this kind. As the purported potential for participation and its threshold for actual participation were by definition latent variables, there was no way to evaluate whether such variations or deviations existed, but if they did, we could have ended up identifying groups that had no substantive meaning. Because we explicitly acknowledged that any groups we identified may not actually exist as taxons in the population or in the 6-Day Sample, however, we could make use of the group solution that appeared to offer the greatest substantive descriptive power. Moreover, over-extraction of groups due to non-normality in the data acts to reduce power to identify group membership covariates (Bauer & Curran, 2003). This possibility made any covariates we identified that much more meaningful.

Because so many participants completed all the young-adult follow-up assessments, we fit the latent class models only to participants missing at least one assessment, and subsequently added the participants who had completed all assessments as an additional group in identifying covariates of participation. The data to which we fit the models was yes/no participation at each assessment, so all variables were dichotomous, meeting that assumption for use of latent class analysis. It was reasonable to consider that the data met the other two assumptions necessary for appropriate use of latent class analysis as well: the model must be identifiable, and the latent class structure can be considered to account for all systematic variance in the data. We discuss post-hoc evaluations of the reasonableness of the latter assumption below. As shown in Table 3, we fit models offering 1-7 groups. Several of the fit indices indicated improved fit with every additional group added to the model, which is common in fitting latent class models, as well as factor analysis. Though there was

some evidence of convergence in the fit statistics, the number of indicated groups was becoming impractically large for our descriptive purposes, and 2 of the 5 tests we used (BIC and the Lo-Mendell-Rubin Adjusted Loss Ratio Test; see Table 3) indicated that 5 groups were appropriate. Under this model, probabilities of latent group membership were high, ranging from .962 to 1.000. This indicated that the model distinguished among the groups quite clearly, corroborating the appropriateness of using it. Each group comprised at least 4% of the sample, suggesting that the model was not unduly over-fitted. Moreover, when we split the sample in two randomly, the fit statistics similarly indicated the 5-class solution, and the two sets of indicated groups had characteristics and relative sizes similar to those in the full sample, The indicated conditional probabilities of participation at each assessment and patterns of lagged correlations among assessment participations also supported the reasonableness of this solution, and these observations also offered post-hoc support for the reasonableness of the assumption underlying latent class analysis that the indicated latent class structure accounted for the systematic variation in the data. We used participants' group assignments from this model, supplemented with one additional group consisting essentially of those who participated in all assessments, for the remainder of our analyses. Two participants were not assigned to groups under the model. These two participants had missed only one assessment but the assessment missed was atypical of those with almost-complete participation. We included them with the full-participation group.

Table 4 shows participation characteristics of the 6 indicated groups. Most groups had quite even sex distributions, but female participation was only 30% in the rather small (66 participants; 6% of total) second group, characterized by especially high participation in the early years, which fell off rapidly in the mid 1950s, so that on average these participants completed only about 6 assessments. Females comprised 57%, however, in the fifth group (62 participants; 5%), which averaged not even 2 assessments. The 1955 assessment was commonly missing in the otherwise highly-participating first group (48 participants; 4%), while the third group (305 participants; 25%) tended to have missed just 1952 and/or 1953. Like the second group, the fourth group (79 participants; 7%) participated highly in the earlier years and then dropped off sharply, but their participation continued high until the early 1960s. Eight participants died during the assessment period; two were in the second group, the rest in the fifth group.

Table 5 shows the covariates of participation group membership we evaluated. Descriptive statistics are shown for each group, along with indications of which covariates distinguished among which groups. In the table, common superscripts indicate statistically indistinguishable means based on analysis of variance with Tukey's-b post-hoc testing to distinguish among groups, and the *p* values shown refer to results of the omnibus ANOVA F test. Tukey's b is a rather conservative post-hoc test of specific group differences, so it is not uncommon to see a significant omnibus test for which it does not identify any specific groups that differ. The groups did not differ in height (despite a significant omnibus test), self-confidence, originality, or desire to excel. All other characteristics distinguished among the groups, though to varying degrees. Some of the distinctions supported commonly observed associations and intuitive and stereotypical notions; some did not. For example, the groups with lowest mood stability, and conscientiousness also reported the most personal problems, and the groups with highest cognitive abilities received more years of schooling. But the groups with highest cognitive abilities had very average father's social class, and the group with lowest cognitive abilities had rather high father's social class. See Table 6.

The personal characteristics gave clear hints as to reasons for the participation patterns. All the high-participation groups (groups 1, 3, and 6) showed rather high perseverance, mood stability, and conscientiousness, reported relatively few personal problems, and faced low health risk. Groups 3 and 6 also had the highest cognitive abilities, also commonly associated with greater tendency to participate in research studies. Though group 1 had the lowest cognitive abilities, it had the highest mood stability and conscientiousness. Group 3, which tended to have missed the 1952 and/or 1953 assessments, had the highest cognitive abilities and most years of schooling; they were 16 and 17 in 1952 and 1953 and may have missed these assessments due to transitions from secondary school at home with parents to university away from home. Group 1, which tended to have missed the 1955 assessment, did receive post-secondary schooling, though likely not university education due to their relatively low cognitive abilities. They were 19 in 1955; this may have coincided with transition from post-secondary school or National Service into new job situations and/or marriage. Likely reasons for group 2 participants to have stopped participating include poor perseverance, mood stability, and conscientiousness, coupled with rather high numbers of years of schooling, job transitions and personal problems, and rather low cognitive abilities. Reasons for poor participation throughout in group 5 were very clear: proportionately these participants faced substantial health risks (6 died,

though some accidentally) and rather many (likely associated) personal problems. Table 6 summarizes these group characteristics.

Participation Patterns at Age-77 Follow-Up

The participation rate at the age-77 follow-up was dramatically different from those in young adulthood. Only 171 (14.2%) of the original 1208 completed the age-77 assessment. The relevant participation rate of 26.8%, however, must be based on those who survived to that age and were invited (636; 52.6% of the original 1208). Several age-11 personal characteristics predicted survival to age 77 and further distinguished between survival and participation at that age. Associations with survival have been addressed in detail elsewhere (Calvin, Batty, Brett, & Deary, submitted; Deary, Batty, Pattie, & Gale, 2008). Participation associations included cognitive abilities, father's social class, perseverance, and years in school. Higher conscientiousness, greater number of moves, and fewer personal problems predicted survival but not participation, and mood stability and height predicted participation but not survival. When both were significantly predictive, the effect sizes of the participant-survivor differences were considerably larger than those of the survivor-deceased differences. Table 7 summarizes reports full results.

Membership in the young-adult participation groups also predicted old-age participation. Table 8 summarizes these results. Groups 3 and 6, which had the highest participation rates in the young-adult assessments, participated disproportionately highly in old age as well. Consistent with difficulty in locating participants during school transitions causing the 1952 and/or 1953 gaps in participation in group 3 rather than gaps in survey administration, members of this group participated in old age at slightly but not significantly greater rates than those in group 6. Groups 2 and 5, which had the lowest participation rates in the young-adult assessments, also were disproportionately likely not to participate in old age. Despite indication that the lower young-adult participation was due to health problems in group 5, this was not because of disproportionately lower survival rate. Group membership did not predict survival at all and early-life participants in group 5 participated in.the old—age follow-up at essentially exactly their expected proportionate rate.

Generalizing Cognitive Ability Associations with Study Participation in Old Age

As noted in introducing this study, cognitive ability is one variable on which participants in research studies tend to show higher performance than population averages. This is typically measured contemporaneously with old age study recruitment by comparing with norm references,

and attributed to current life circumstances. This study and a few others have indicated, however, that old-age selection patterns for cognitive ability may reflect its long-term stability (Deary, 2014) rather than anything in particular about current life circumstances. When long-term stability is under consideration, differences in administered tests can make it especially difficult to obtain the needed to make valid assessments of extent of selection on cognitive ability. Such comparison was possible here, but we were interested not just in understanding the degree of selection in this sample, but to what degree we could generalize beyond this sample about likely rates of selection on cognitive ability in other old-age samples. For this, we needed other old-age samples that had taken the same cognitive test.. Fortunately, several such samples were available.

All were based on the SMS surveys of 1932 and 1947, which limits generalizability to some degree. But all were recruited originally for extensive study only in old age, unlike the 6-Day Sample, and completely independently, at different times and at different participant ages. Some participants were born in 1936 as were the 6-Day Sample participants, but others were born in 1921, so they did not all grow up in the same generation. In particular, the Aberdeen and Lothian Birth Cohorts differed at least in age, method of recruitment, and mean age-11 MHT scores. Table 9 shows results of these comparisons. All scores shown were obtained when participants were age 11. Together they indicate very similar degrees of selection on cognitive ability: mean scores were considerably and similarly above the full sample averages, and standard deviations were similar. This suggests that the 27% participation rate in the 6-Day Follow-Up reflected the highly population-representative nature of the recruitment process, and many aging studies with higher apparent participation rates may have similar participation rates relative to their underlying populations. Skews were consistently negative, even at age 11. Negative age-11 skews in the SMS population-level scores indicate that this is a characteristic of the MHT, the populations, or both, but the greater negative skews in the old age distributions reflected the sample selection that took place on cognitive ability: some people with lower cognitive ability at age 11 did participate in old age, but they were less likely to do so than those of higher cognitive ability. This is shown graphically in Figure 1. The figure also shows proportions of those surviving to age 77 who participated. Greater proportions of those with higher age-11 cognitive ability scores survived to age 77, but among those who survived, greater proportions of those with higher age-11 scores participated. Point biserial correlations between age-11 cognitive ability scores and survival to age 77 and participation at that age were .17 and .28, respectively. The large

correlations between age-11 and old-age cognitive ability in all the samples suggested that the same patterns were present in contemporaneous cognitive abilities in the full surviving population eligible for participation. This indicates that the extent of selection for cognitive ability in the follow-up study of the Six-Day sample should be considered rather typical of old-age samples recruited by mail.

DISCUSSION

We made use of an unusually population-representative sample assessed originally at age 11 and followed extensively throughout young adulthood and then again in old age to explore selection effects of cognitive ability, personality, and other personal characteristics on study participation. Participation rates were much higher in young adulthood: over half the original sample completed all 13 follow-up assessments. We used latent class analysis to identify groupings of participants within the data that tended to have co-occurring characteristics that were associated with patterns of willingness and availability to participate in the study over time. We did not consider these groupings likely to represent robust taxons, but instead 'fuzzy sets' that offered descriptive power. The 6 groupings of the data indicated by these analyses suggested that there were specific reasons such as physical moves and/or transitions to new life situations such as new jobs or educational programmes for many of the relatively isolated missed assessments. More chronic missed participation, however, was associated with lower cognitive abilities and father social class, as well as teacher ratings of childhood personality, particularly perseverance, mood stability, and conscientiousness, and the presence of chronic health and personal problems during young adulthood. Despite these general patterns, which have also been noted in many prior studies (Menard, 2002, Taris, 2000), our latent class analysis suggested several more specific patterns as well. One small group (group 1) with strong participation had the lowest cognitive ability, but, contrary to the overall patttern of correlation between cognitive ability and higher father social class (.22), this group had rather high father social class. In addition, they had the highest conscientiousness and mood stability. Another small group also had rather low cognitive ability and rather high father social class, but this group had very poor participation, which could be attributed to chronic health and personal problems, as well as to rather (likely often associated) low perseverance, mood stability, and conscientiousness.

In contrast to the high participation in the young adult assessments, participation at the age-77 follow-up was quite low. Some original participants of course were not even invited to participate due to death, emigration from the UK, or inability to locate, but many of those invited either did not

reply or refused to participate as well. This likely can be attributed in part to the recruitment methods: participants were recruited personally in childhood, and visited in their homes for assessment during young adulthood. The annual assessments likely fostered senses of continuity and even possibly community. By age 77, however, these would have been lost, and the sample members received a postal recruitment, including a box of materials to complete and in which to return a self-assessment, which may have seemed a sharp contrast to prior methods. Completion involved some intellectual complexity and required some commitment of indivdual time and effort, and even special pick-up at local post offices in some cases. Nonetheless, such a recruitment and assessment method is not uncommon in old-age surveys, and the degree of sample selection here was very comparable to those in the other old-age samples to which we could make direct comparisons. These other samples, though also recruited by mail, had been assessed in person. This suggests that our selectivity results likely can be generalized at least to old-age samples recruited by mail, By way of comparson, there were no differences in participation in the 1951-63 assessments among those who emigrated, did not reply to the recruitment invitation, or refused to participate at age 77. Among the variables we considered, these groups differed only in Terman-Merrill IQ and years of education (MHT Test scores showed the same pattern but differences were not significant), with emigrants having the highest mean IQ and years of education (109.5 IQ, 2.0 tertiary education), then those who refused (101.7, 1.4), then those who did not reply (99.8, 1.1). Only the emigrants were statistically distinguishable (effect size ~.45 in both cases).

The overall tendencies for early-life cognitive abilities, years in school, and even the teacher-rated personality characteristics of perseverance and mood stability to predict participation were stronger in old age than they had been in youth, but the health and personal problems during young adulthood that had been associated with lack of participation at the time were not associated with participation in old age. Because young adult health risk was not associated with survival to age 77, this suggested that many of those problems had been overcome. The relative rates of partipation in old age among the participation groups we identified in young adulthood directly paralleled the young adult rates, with one exception: the rank ordering of the groups by participation rate in young adulthood was 5, 2, 4, 3, 1, 6 while that in old age was 5, 2, 4, 1, 6, 3. This appeared to be because, in young adulthood, those with the highest cognitive abilities also completed the most years of school and some kind of school transition caused some of them to miss assessment at age 16 and/or 17.

Such transitions presumably did not preclude participation in old age. These associations did not arise from differential survival by group.

Ability of early-life characteristics to predict old-age participation in the 6-Day Sample was strong. This suggests that the common tendency to attribute sampling bias in epidemiological and psychological aging studies to contemporaneous circumstances is an oversimplification of the processes involved. The actual processes likely involve lifetime-stable characteristics that also contribute to emergence of the whole constellation of circumstances in which people spend their later lives. For example, people of higher cognitive abilities and more perseverant, stable, and conscientious personalities also obtained more education in this sample (as they do more generally), and this education may have fostered greater interest in scientific research in health that has guided lifestyle choices throughout life. But greater cognitive abilities also tend to make the prospect of completing a long assessment involving questionnaires and even use of unfamiliar equipment for physical tests less daunting, and this general can-do capacity also likely serves people well in addressing challenges throughout life. This makes inferences of purely environmental causal associations much more difficult. It has been noted before with respect to cognitive ability and middle-age participation (Nishiwaki, et al., 2005), but it seems particularly remarkable that it also applied here at age 77 to personality characteristics, independent of cognitive abilities, rated by teachers at age 11. To illustrate this apparently previously unobserved personality association, we constructed a personality composite variable based on perseverance, mood stability, conscientiousness, and desire to excel, which loaded together on the first factor in factor analysis, and constructed ranges of this variable directly analogous in dispersion from the mean to the IQ ranges shown in Figure 1 for cognitive abilities. Results are shown in Figure 2.

The primary limitation of this study was that its results may be specific to the historical context in which the 6-Day Sample was originally collected, and/or the particular recruitment and assessment methods used in the young-adult and old-age follow-ups. The Scottish Mental Surveys were essentially unique in history anywhere, and they took place within the particular context of Scottish history at that time. Moreover, all the samples we obtained to make comparisons of degree of selection were also related to the Scottish Mental Surveys, and even the Nishiwaki et al., (2005) study that found very similar selection patterns involving childhood cognitive abilities was conducted in Scotland as well. Still, the other samples were all collected independently using different recruiting

methods, at different ages, from different areas, and from more than a single birth cohort. In the Nishiwaki et al. (2005) study, different measures were used as well. The results we could compare were highly consistent, indicating that our results should not be considered at all unique, and suggesting that they may actually apply quite broadly. Other, likely smaller, limitations include the possibility of biases due to associations between cognitive abilities and/or other characteristics and inability to locate participants in old age, or propensity to have emigrated.

Conclusion

In a cohort for which clear data on extent of population representativeness were available, study participation in old age was much lower than in young adulthood, sharply reducing the degree to which the sample represented the population surviving to old age recruitment. Age-11 cognitive abilities and even teacher-rated personality characteristics related to conscientiousness or dependability were associated with study participation at age 77, more strongly than with participation at ages much closer to original recruitment. Organizers of aging studies need to be aware that their samples are likely highly selected relative to the populations from which they are drawn, and that the characteristics on which they are selected have likely been stable throughout the lives of their participants, as well as those in the population who are not participating. These lifetime-stable characteristics have probably done much along the way to shape the circumstances in which both participants and non-participants live in old age. This means research attention the processes involved in this, as well as greater circumspection in attributing research results to current environmental circumstances than is often shown.

ACKNOWLEDGEMENTS

The 6-Day Sample follow-up study is supported by the UK Medical Research Council's contributions to the cross-council Life Long Health and Wellbeing Initiative (MRC G1001401/1). The work was undertaken by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross-council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the Biotechnology and Biological Sciences Research Council (BBSRC) and Medical Research Council (MRC) is gratefully acknowledged.

REFERENCES

Akaike, H. (1983). Information measures and model selection. *Bulletin of the International Statistics Institute, 50*, 277-290.

Bauer, D. J., & Curran, P. J. (2003). Distributional assumptions of growth mixture models: Implications for overextraction of latent trajectory classes. *Psychological Methods, 8*, 338-363.

Beaver, K. M. (2013). Intelligence and selective attrition in a nationally representative and longitudinal sample of Americans. *Personality and Individual Differences, 55*, 157-161.

Boutwell, B. B., Beaver, K. M., & Barnes, J. C. (2012). Role of breastfeeding in childhood cognitive development: A propensity score matching analysis. *Journal of Pediatrics and Child Health, 48,* 840-845.

Brett, C. E., & Deary, I. J. (2014). Realising health data linkage from a researcher's perspective: Following up the 6-Day Sample of the Scottish Mental Survey 1947. *Longitudinal and Life Course Studies, 5,* 283-298.

Calvin, C. M., Batty, G. D., Brett, C. E., & Deary, I. J. (submitted). Socially engaged children live longer: A 65-year follow-up of a population-representative sample in Scotland. *Psychosomatic Medicine.*

Calvin, C. M., Deary, I. J., Fenton, C., Roberts, B., Der, G., Leckenby, N., et al. (2011). Intelligence in youth and all-cause mortality: Systematic review with meta-analysis. *International Journal of Epidemiology, 40*, 626-644.

Cooney, T. M., Schaie, K. W., & Willis, S. L. (1998). The relationship between prior functioning on cognitive and personality dimensions and subject attrition in longitudinal research. *Journals of Gerontology: Psychological Sciences Series B, 43*, 12-17.

Deary, I. J. (2014). The stability of intelligence from childhood to old age. *Current Directions in Psychological Science*, *23*, 239-245.

Deary, I. J., Batty, G. D., Pattie, A., & Gale, C. G. (2008). More intelligent, more dependable children live longer: A 55-year longitudinal study of a representative sample of the Scottish Nation. *Psychological Science*, *19*, 874-880.

Dykiert, D., Gale, C. R., & Deary, I. J. (2009). Are apparent sex differences in mean IQ scores created in part by sample restriction and increased male variance? *Intelligence, 37*, 42-47.

Hunt, E., & Madhyastha, T. (2008). Recruitment modeling: An analysis and an application to the study of male-female differences in intelligence. *Intelligence, 36,* 353-363.

Johnson, W., Brett, C., & Deary, I. J. (2010). Intergenerational class mobility in Britain: A comparative look across three generations. *Intelligence, 38*, 268-281.

Johnson, W., McGue, M., & Deary, I. J. (2013). Normative cognitive aging. In D. Finkel & C. A. Reynolds (Eds.). *Behavior Genetics of Cognition* (pp. 135-168). New York: Springer.

Kerr, M., Lambert, W. W., & Bem, D. J. (1996). Life course sequelae of childhood shyness in Sweden: Comparison with the United States. *Developmental Psychology, 32*, 1100-1105.

Knight, R. (1967). Changes in the occupational structure of the working population. *Journal of the Royal Statistical Society - Series A, 130*, 408-422.

Lo, Y., Mendell, N. R., & Rubin, D. B. (2001). Testing the number of components in a normal mixture. *Biometrika, 88*, 767-778.

De Luna, X., & Lundin, M. (2014). Sensitivity analysis of the unconfoundedness assumption with an application to an evaluation of college choice effects on earnings. *Journal of Applied Statistics, 41,* 1767-1784.

Lynam, D., Moffitt, T. M., & Stouthhammer-Loeber, M. (1993). Explaining the relation between IQ and deliquency: Class, race, test motivation, school failure, or self-control? *Journal of Abnormal Psychology, 102*, 187-196.

MacPherson, J. S. (1958). *Eleven-Year-Olds Grow Up.* London: University of London Press.

Maxwell, J. (1969). *Sixteen Years On: A Follow-Up of the 1947 Scottish Survey.* London: University of London Press.

McLachlan, G., & Peel, D. (2000). *Finite Mixture Models.* New York: John Wiley & Sons.

Menard, S. (2002). *Longitudinal Research, 2nd Ed.* Thousand Oaks, CA: Sage.

Murray, A. L., Johnson, W., McGue. M., & Iacono, W. G. (2014). How are intelligence and personality related to one another? A re-examination of the intelligence compensation hypothesis. *Personality and Individual Differences, 17,* 17-53.

Nagin, D. S., & Tremblay, R. E. (2005). Latent trajectory groups: Fact or a useful statistical fiction. *Criminology, 43,* 873-904.

Nishiwaki, Y., Clark, H., Morton, S. M., & Leon, D. A. (2005). Early life factors, childhood cognition and postal questionnaire response rate in middle age: The Aberdeen Children of the 1950s study. *BMC Medical Research Methodology, 5*, doi:10.1186/1471-2288-5-16.

Paterson, L., Pattie, A., & Deary, I. J. (2011). Social class, gender and secondary education in Scotland in the 1950s. *Oxford Review of Education, 37*, 383-411.

Rabbitt, P., Lunn, M., & Wong, D. (2008). Death, dropout, and longitudinal measurements of cognitive change in old age. *Journals of Gerontology. Series B, Psychological Sciences and Social Sciences, 63,* 271–278.

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methods, 25*, 111-163.

Roth, P. L., Le, H., Oh, I. S., Van Iddekinge, C. H., Buster, M. A., Robbins, S. B., & Campion, M. A. (2014). Differential validity for cognitive ability tests in employment and educational settings: Not much more than range restriction? *Journal of Applied Psychology, 99,* 1-20.

Sclove, L. S. (1987). Application of model selection criteria to some problems in multivariate analysis. *Psychometrika, 52*, 333-343.

Scottish Council for Research in Education. (1933). *The Intelligence of Scottish Children.* London: University of London Press.

Scottish Council for Research in Education. (1949). *The Trend of Scottish Intelligence.* London: University of London Press.

Scottish Council for Research in Education. (1953). *Social Implications of the 1947 Scottish Mental Survey.* London: University of London Press.

Sliwinski, M., Hoffman, L., & Hofer, S. M. (2010). Evaluating convergence of within-person change and between-person age differences in age-heterogeneous longitudal studies. *Research in Human Development, 7*, 45-60.

Stang, A. (2003). Nonresponse research - An underdeveloped field in epidemiology. *European Journal of Epidemiology, 18*, 929-931.

Taris, T. W. (2000). *A Primer in Longitudinal Data Analysis.* Thousand Oaks, CA: Sage.

United Nations. (2009). *World Population Aging 2009.* New York: UN Department of Economic and Social Affairs, Population Division.

Table 1
The 6-Day Sample and the Scottish Mental Survey

	6-Day Sample, n=1208			Scottish Mental Survey, n=70,805			Effect Size of Mean
	Mean	SD	Skew	Mean	SD	Skew	Difference
Full Sample							
Age at MHT	10.98	.33	.27	10.93	.29	-.08	.17
% Female	51.16	---	---	49.45	---	---	1.71
Family Position	2.41	1.77	1.83	2.51	1.81	1.62	-.06
Family Number	3.72	2.23	1.16	3.78	2.22	1.31	-.03
MHT Score	37.43	15.76	-0.31	36.64	15.77	-.34	.05
Females							
Age at MHT	10.96	.32	.19	10.93	.29	-.07	.10
Family Position	2.41	1.84	1.87	2.54	1.84	1.58	-.07
Family Number	3.65	2.21	1.24	3.81	2.25	1.46	-.07
MHT Score	37.76	14.86	-.27	37.46	15.06	-.36	.02
Males							
Age at MHT	11.00	.35	.31	10.93	.29	-.08	.24
Family Position	2.41	1.70	1.77	2.48	1.79	1.65	-.04
Family Number	3.80	2.26	1.08	3.75	2.18	1.15	.02
MHT Score	37.07	16.67	-.32	35.83	16.40	-.30	.08

Note: SD is standard deviation. MHT is Moray House Test. Effect size is
(6-Day mean-SMS mean)/SMS SD, except for sex, which is simple percentage difference.

Table 2
Follow-Up Participation 1951-63, Ages 15-27, and 2013, Age 77

By Year			By Number of Occasions		
Year	Participants	Percent	Number of Occasions	Participants	Percent
1951	1162	96.2	13	641	53.1
1952	980	81.1	12	188	15.6
1953	795	65.8	11	157	13.0
1954	1122	92.9	10	50	4.1
1955	1074	88.9	9	30	2.5
1956	1124	93.0	8	19	1.6
1957	1115	92.3	7	23	1.9
1958	1088	90.1	6	17	1.4
1959	1078	89.2	5	12	1.0
1960	1060	87.7	4	7	0.6
1961	1021	84.5	3	13	1.1
1962	1000	82.8	2	24	2.0
1963	1070	88.6	1	13	1.1
2013	171	14.2	0	14	1.2

Table 3
Fit Statistics for Latent Class Participation Models

Number of Classes	-2 *Log likelihood (df)	AIC	BIC	Sample-Size Adjusted BIC	Parametric Bootstrapped LRT Test	Lo-Mendell-Rubin Adj. LRT Test
1	3845.2 (13)	7716.4	7772.6	7731.4	----	----
2	2560.0 (27)	5173.7	5290.6	5204.9	2570.65 (.0000)	2541.95 (.0000)
3	2219.3 (41)	4520.5	4698.0	4567.8	681.17 (.0000)	673.57 (.0000)
4	2087.0 (55)	4283.9	4522.0	4347.4	264.62 (.0000)	261.67 (.0000)
5	2402.6 (69)	4177.8	4476.4	4257.4	134.14 (.0000)	132.64 (.0000)
6	1983.2 (83)	4132.4	4491.6	4228.1	73.38 (.0572)	72.56 (.0588)
7	1956.9 (97)	4107.7	4527.5	4219.6	52.71 (.0030)	52.12 (.0032)

Note: Models were fit using only the participants who missed at least one assessment. LRT is Loss Ratio Test; all had 14 df. For both tests, p-values indicate probability that one fewer number of groups would be more appropriate, with .05 being a typical threshold for significance.

Table 4
Participation Characteristics of Latent Early-Life Participation Groups

Group	% Female	Mean (SD) Number of Assessments	Number in Group (%)	Participation Trends
1	51	11.9 (.2)	51 (4.2)	Missing 1-3 randomly, but especially 1955
2	31	6.3 (1.5)	65 (5.4)	High in early years, falls off in mid 1950s
3	48	11.3 (.8)	316 (26.2)	Missing mostly just 1952 and/or 1953
4	48	9.2 (1.3)	65 (5.4)	High in early years, falls off in early 1960s
5	56	1.6 (1.1)	63 (5.2)	Poor throughout
6	53	13.0 (.1)	648 (53.6)	Essentially complete
			1208	

Note: Latent class models were fit using only participants who had missed at least one assessment. We selected the 5-class solution. The sixth group consisted of those who completed all assessments, plus 7 whom the program could not assign to groups. All 7 had missed only the 1963 assessment.

Table 5
Descriptive Statistics for Personal Characteristics of Participation Groups: Mean (SD)

	Group						
	1	2	3	4	5	6	Total
Moray House Test	30.7[a]	30.8[a]	38.7[b]	31.5[a]	34.0[a,b]	38.9[b]	37.4
p <.001	(13.6)	(15.6)	(16.7)	(15.4)	(14.5)	(15.2)	(15.8)
Terman-Merrill IQ	93.5[a]	96.4[a,b]	106.0[c]	97.0[a,b]	97.5[a,b]	103.3[b,c]	102.6
p <.001	(13.8)	(16.1)	(21.0)	(15.9)	(17.0)	(20.6)	(20.1)
Father's Social Class	3.12[a,b]	3.69[b,c]	3.31[a,b]	3.80[b,c]	3.19[a]	3.35[a,b]	3.37
p=.001	(0.82)	(0.92)	(1.00)	(1.09)	(1.02)	(0.93)	(0.97)
Height	53.18	53.40	54.04	53.32	54.33	54.16	54.01
p=.020	(2.90)	(2.80)	(2.96)	(2.50)	(2.73)	(2.82)	(2.85)
Self-Confidence	-.18	-.03	.03	.01	-.14	.02	.00
p=.386	(.71)	(.78)	(.75)	(.97)	(.97)	(.76)	(1.00)
Perseverance	.03[a,b]	-.33[b]	-.02[a]	.02[a,b]	-.09[a,b]	.05[a]	.00
p=.011	(.64)	(.84)	(.76)	(.94)	(.87)	(.79)	(1.00)
Mood Stability	.05[a,c]	-.43[b]	.03[a]	-.26[b,c]	-.28[b,c]	.07[a,c]	.00
p <.001	(.78)	(.69)	(.88)	(.96)	(.94)	(.85)	(1.00)
Conscientiousness	.07[a]	-.40[b]	.02[a]	-.05[a,b]	-.15[a,b]	.04[a]	.00
p=.003	(.85)	(.82)	(.87)	(.86)	(1.00)	(.86)	(1.00)
Originality	.02	-.07	-.03	-.10	-.05	.03	.00
p=.550	(.77)	(.85)	(.73)	(.69)	(.73)	(.69)	(1.00)
Desire to Excel	.03	-.04	-.05	.12	-.06	.02	.00
p=.672	(.84)	(.85)	(.83)	(.93)	(.81)	(.83)	(1.00)
Number of Moves	-.01[a]	-.04[a]	.00[a]	-.01[a]	-.49[b]	.00[a]	.00
p <.001	(.07)	(.16)	(.08)	(.10)	(.59)	(.08)	(.11)
Number of Jobs	-.00[a]	-.04[a]	.00[a]	.06[b]	-.49[c]	.00[a]	-.02
p <.001	(.08)	(.16)	(.08)	(.14)	(.45)	(.08)	(.16)
Health Risk	.00[a]	.03[a]	.00[a]	-.02[a]	.74[b]	.00[a]	.01
p <.001	(.09)	(.20)	(.08)	(.08)	(.87)	(.08)	(.14)
Personal Problems	.01[a]	.04[a,b]	.13[a]	.04[b]	.07[c]	.01[a,b]	.02
p <.001	(.04)	(.11)	(.04)	(.07)	(.20)	(.04)	(.06)
Years in School	.08[a]	.11[a,b]	.17[b]	.07[a]	.05[a]	.12[a,b]	.12
p <.001	(.13)	(.19)	(.20)	(.12)	(.13)	(.15)	(.17)

Note: Personality assessments were completed by teachers and correlated with intelligence scores. Means shown here were residuals after partialling Terman-Merrill IQ. Larger numbers indicate lower father's social class. Reports of residence and job moves, health risk, and personal problems were standardized per assessment. Years in school were per assessment. Where there were significant mean differences, measures with statistically indistinguishable means have the same superscripts. Measures that showed no significant mean differences have no superscripts. We did not adjust signficance levels for multiple testing across the personal characteristics as the analyses were primarily explorative. See Table 4 for description of groups. Given *p* values are for omnibus ANOVA F test.

Table 6
Distinctive Characteristics of Participation Groups

Group	Distinctive Characteristics
1 (Dedicated)	Lowest cognitive ability; highest father social class; lowest self-confidence*; highest conscientiousness and high mood stability; high participation, but likely to have missed 1955; smallest group.
2 (Drop-outs)	Low cognitive ability and father social class; lowest perseverance, mood stability, and conscientiousness; above average personal problems; high particiaption in early years, but dropped off rapidly in mid-1950s; 70% male.
3 (Capable)	Highest cognitive ability, self-confidence*, and number of years of schooling; high participation but likely to have missed 1952 and/or 1953; largest group with any missed assessments.
4 (Edgy)	Rather low cognitive ability and mood stability; lowest father social class; highest desire to excel* and number of moves*; most personal problems; relatively few years in schooling; high participation until 1960s.
5 (Troubled)	Low cognitive ability; high father social class; lowest self-confidence*, perseverance; low mood stability, conscientiousness; lowest numbers of moves, jobs; highest health risk and personal problems; lowest number of years in school; most of those who died in this group; 56% female; participation poor throughout.
6 (Stable)	Rather high cognitive ability; highest self-confidence*, perseverance, mood stability, conscientiousness; relatively many years in school; participated throughout; by far largest group.

Note: *No mean differences were significant.

Table 7
Early-Life Descriptives of Age-77 Participation Status Groups

	Mean (SD)		
	Full 6-Day Sample n=1208	Alive at Age 77 n=791	Participating at Age 77 n=171
Moray House Test	37.4 (15.8)	39.4 (15.3)*	48.1 (11.6)*
Terman-Merrill IQ	102.6 (20.1)	105.1 (20.6)*	115.6 (19.7)*
Father's Social Class	3.4 (1.0)	3.3 (1.0)*	3.1 (1.0)*
Height	54.0 (2.8)	54.1 (2.9)	54.7 (2.8)*
Self-Confidence	0.0 (1.0)	-.01 (1.00)	.04 (.98)
Perseverance	0.0 (1.0)	.06 (.99)*	.19 (.94)*
Mood Stability	0.0 (1.0)	.02 (.97)	.20 (1.11)*
Conscientiousness	0.0 (1.0)	.06 (.97)*	.13 (1.00)
Originality	0.0 (1.0)	.01 (.97)	.10 (1.01)
Desire to Excel	0.0 (1.0)	.03 (1.00)	.10 (.95)
Number of Moves	.00 (.11)	.00 (.10)*	.00 (.08)
Number of Jobs	-.02 (.16)	-.02 (.18)	-.01 (.07)
Health Risk	.01 (.14)	.00 (.11)	.01 (.10)
Personal Problems	.02 (.06)	.01 (.04)*	.01 (.04)
Years in School	.12 (.17)	.14 (.18)*	.23 (.20)*

Note: All measures in youth. Personality assessments were completed by teachers and correlated with intelligence scores. Means shown here were residuals after partialling Terman-Merrill IQ. Larger numbers indicate lower father's social class. Asterisks in the Alive column indicate mean differences between survivors and deaths; those in the Participating column indicate mean differences between participants and survivors, at $p<.01$ to adjust for multiple testing.

Table 8
Relative Survival and Participation Rates by Group

	Group						
	1	2	3	4	5	6	Total
	Dedicated	Drop-outs	Capable	Edgy	Troubled	Stable	
Survivors							
Actual	30	38	218	34	41	430	791
Expected	33.4	42.6	206.9	42.6	41.3	424.3	791
Relative Rate	.90	.89	1.05	.80	.99	1.01	
Participants							
Actual	3	2	55	3	1	107	171
Expected	7.2	9.2	44.7	9.2	8.9	91.7	171
Relative Rate	.42	.22	1.23	.33	.11	1.17	

Note: Relative rate is ratio of actual to expected participation. Differences were significant for participants but not for survivors. See Table 4 for descriptions of groups.

Table 9
Descriptive Statistics for Age-11 Moray House Test in Comparable Samples

	N	Mean	Standard Deviation	Skew	Correla-tion
Scottish Mental Survey 1947	64,532	36.6	15.8	-.34	---
Scottish Mental Survey 1932	81,140	34.0	15.4	-.20	---
Full 6-Day Sample	1,112	37.4	15.8	-.31	---
6-Day Sample Alive at 77	732	39.4	15.3	-.38	---
6-Day Sample Follow-Up Participants	159	48.2	11.6	-.78	.61
Lothian Birth Cohort 1936 , recruited age 70	1,028	49.0	11.8	-.78	.67
Lothian Birth Cohort 1936, followed up age 73	816	49.5	12.0	-.90	.62
Lothian Birth Cohort 1921, recruited age 79	550	46.4	12.0	-.49	.66
Lothian Birth Cohort 1921, followed up age 87	173	48.4	11.2	-.42	.51
Lothian Birth Cohort 1921, followed up age 90	114	49.0	10.8	-.47	.54
Aberdeen Birth Cohort 1936, recruited age 64	353	43.0	12.7	-.47	.65
Aberdeen Birth Cohort 1921, recruited age 77	77	43.1	12.7	N/A	.63

Note: Correlations between age-11 and later Moray House Test scores were taken from the papers published on the various follow-up samples of the Scottish Mental Survey at follow-up. 6-Day Sample and Aberdeen Birth Cohort 1921 correlations were with g-factor scores of the tests administered, as was Lothian Birth Cohort 1936 at age 73.

FIGURE CAPTIONS

Figure 1 – 6-Day Sample Follow-Up participation rates at age 77 by level of age-11 Terman-Meriill IQ. 'Full' refers to the original 6-Day Sample.

Figure 2 – 6-Day Sample Follow-Up participation rates by level of a personality composite of age-11 teacher-rated persistence, mood stability, conscientiousness, and desire to excel, independent of IQ. 'Full' refers to the original 6-Day Sample.

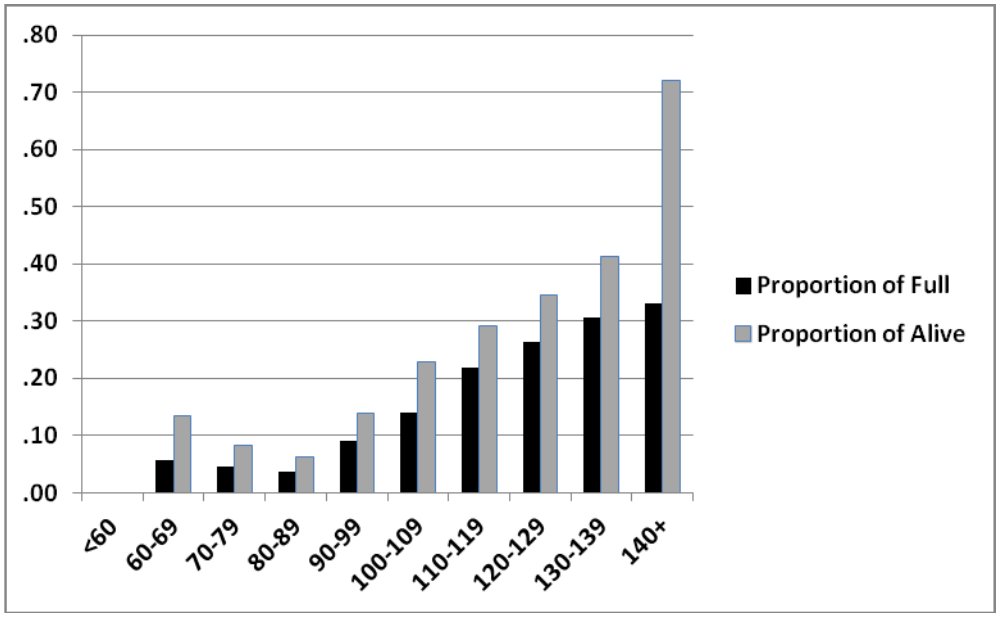

Figure 1

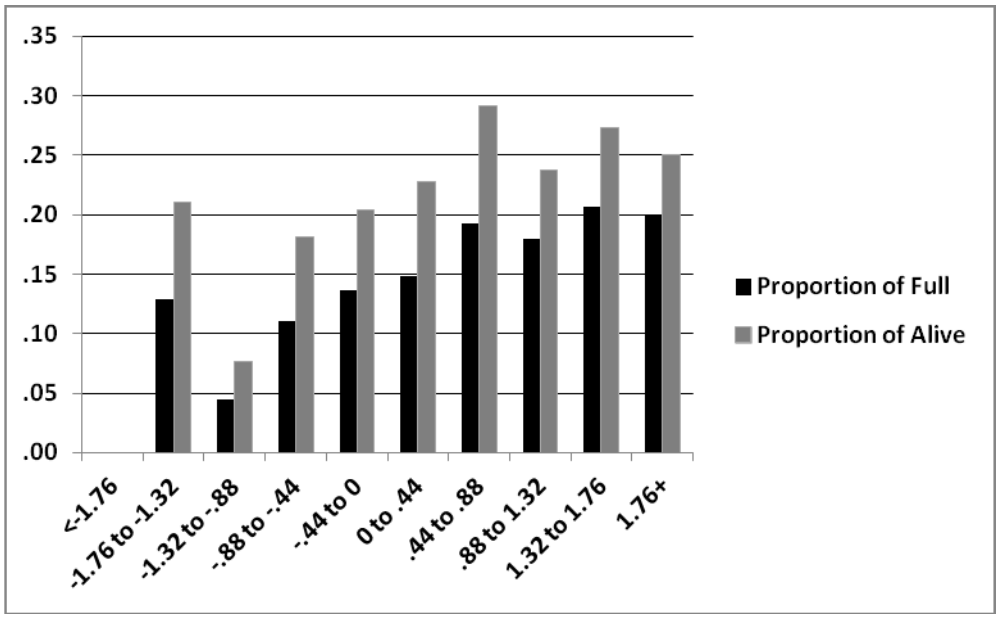

Figure 2